IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

Esophageal Abnormality detection using DenseNet based Faster R-CNN with Gabor features

NOHA GHATWARY[1,2], XUJIONG YE[1], AND MASSOUD ZOLGHARNI[3]

[1]Computer Science Department, University of Lincoln, Lincoln, United Kingdom
[2]Computer Engineering Department, Arab Academy for Science Technology & Maritime Transport, Alexandria, Egypt
[3]School of Computing and Engineering, University of West London, London, United Kingdom

Corresponding author: Noha Ghatwary (e-mail: nghatwary@lincoln.ac.uk, noha.ghatwary@aast.edu).

ABSTRACT Early detection of esophageal abnormalities can help in preventing the progression of the disease into later stages. During esophagus examination, abnormalities are often overlooked due to the irregular shape, variable size and the complex surrounding area which requires a significant effort and experience. In this paper, a novel deep learning model which is based on Faster Region-Based Convolution Neural Network (Faster R-CNN) is presented to automatically detect abnormalities in the esophagus from endoscopic images. The proposed detection system is based on a combination of Gabor handcrafted features with CNN features. The Densely Connected Convolution Networks (DenseNets) architecture is embraced to extract CNN features providing a strengthened feature propagation between the layers and allay the vanishing gradient problem. To address the challenges of detecting abnormal complex regions, we propose fusing extracted Gabor features with CNN features through concatenation to enhance texture details in the detection stage. Our newly designed architecture is validated on two datasets (*Kvasir and MICCAI 2015*). Regarding the Kvasir, the results show an outstanding performance with a recall of 90.2% and precision of 92.1% with a mean of average precision (mAP) of 75.9%. While for the Miccai 2015 dataset, the model is able to surpass the state-of-the-art performance with 95% recall and 91% precision with mAP value of 84%. Experimental results demonstrate that the system is able to detect abnormalities in endoscopic images with good performance without any human intervention.

INDEX TERMS Detection, DenseNet, EAC, Esophagitis, Faster R-CNN, HD-WLE.

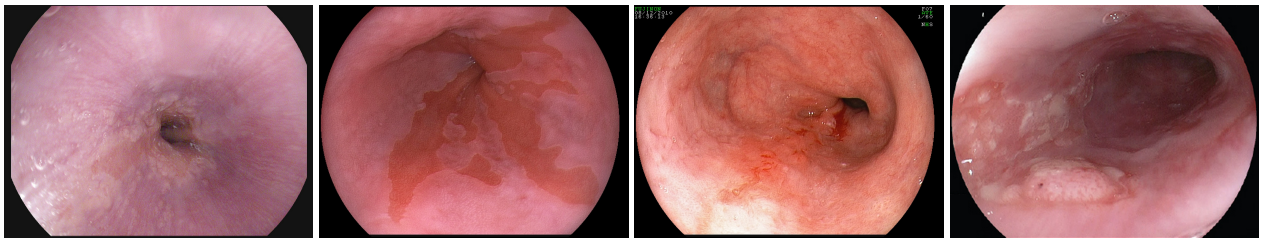
I. INTRODUCTION

Esophageal cancer (EC) is the 7th most common cancer in adults worldwide [1] with a low survival rate on a 5-year plan [2]. EC usually occurs in the cells that fill inside of the esophagus and can appear anywhere along the esophagus tube. It is classified according to the type of cells (gland or squamous) into Esophageal Adenocarcinoma (EAC) and Squamous Cell Carcinoma (SCC) [3]. Early esophageal cancer typically causes no symptoms and mainly arises from untreated/unmonitored premalignant abnormalities. Any inflammation or a small change in the cells of the esophagus tube is considered as a precancerous stage such as Esophagitis and Barrett's Esophagus (BE). *Esophagitis* is an inflammation of the lining of the esophagus that may develop into BE [4]. It usually occurs when either an infection or irritation occurs in the esophagus tube. *BE* is the change of the normal cells with metaplastic intestinal epithelium [5]. BE is considered the main precancerous condition affecting the lower region esophagus tube. The detection and treatment of esophageal abnormalities (precancerous and early cancer stages) are essential as it can increase the survival rate from 19% to 80% [6].

Different endoscopy tools can be used to examine the gastrointestinal tract where the esophagus is located, the High-Definition White Light Endoscopy (HD-WLE) and WLE are considered the most used tools for examination to detect abnormalities in the esophagus. The process of detection is challenging as abnormalities (including early cancer stages) can be located randomly throughout the esophagus tube with various sizes and appearances which makes it difficult to capture by unexperienced endoscopists [7]. Fig. 1 illustrates examples from endoscopic images capturing different types of abnormalities (Esophagitis, BE, EAC & SCC).

Computer Aided Detection (CAD) systems have been

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access



(a) Esophagitis (b) BE (c) EAC (d) SCC

FIGURE 1: Example of the endoscopic view for the four different abnormality types: (a) Esophagitis, (b) BE, (c) Esophageal Adenocarcinoma (EAC), (d) Squamous cell carcinoma (SCC)

developed to assist physicians as a second opinion by extracting features from medical images to automatically detect abnormalities. CAD systems that support the analysis of esophageal abnormality have started to grab more attention with the increase of the number of patients. In previous studies [8], [9], handcrafted features such as *color, texture* and *shape* were extracted from endoscopic images and used in CAD models to find abnormalities. The selection of the appropriate handcrafted features is challenging as it should be chosen according to the characteristics of the image in each application. Lately, deep learning has been widely applied in medical image detection and classification field by extracting features through convolutional neural networks (CNNs) [10]. The CNN deep networks are able to generate features from the images through learning from the dataset, increasing its generalization and scalability for automatic detection [11]. The standard CNN architecture for feature extraction is composed of a series of convolutional filters with reduction layers [12].

In literature, different CNN architectures are constructed to learn and provide informative features for the detection and classification methods such as: (*AlexNet [13], VGG'16 [14], ResNets [15], etc...*). The depth of the CNN network shows a significant impact on the performance of the network but getting deeper without changing in the structure can lead to poor performance, loss of information and facing vanishing the gradient parameter [16]. To overcome these problems, Huang *et al.* [17] introduced the Densely Connected Convolution Networks (DenseNet). The advantages of DenseNet architecture is that it lowers the number of parameters, improves the gradient and information flow throughout the network which makes it easier to train. Also, DenseNet encourages feature reuse by connecting the output of each layer to another layer.

Recently, the combination of handcrafted features with CNN features showed that it can boost the performance of the model [18]. Texture features such as Gabor features has shown its effectiveness when merged with CNN features by providing low-level texture information [19]. The advantage of merging both sets of features have been confirmed in different studies [20]–[23]. Gabor filters has been known for strengthening the texture details provided through spatial information. Additionally, concerning the esophageal abnormality detection, the Gabor features have shown its efficiency in detecting the intestinal juices [24].

There exists various object detection methods that rely on CNN features for final detection, including Regional-Based Convolution Neural Network (R-CNN) [25], Fast R-CNN [26] and Faster R-CNN [27]. The R-CNN generates region proposals by using *selective search algorithm*, then CNN features are extracted from each proposal and classified using support vector machine (SVM). The overhead of applying CNN to each proposal caused the method to be too slow. The Fast R-CNN solved this problem by applying the selective search on the CNN feature map generated from input image. Also, a region-of-interest pooling (ROI pool) layer has been added to the end of the network to classify the features of proposals using softmax. The time consumed for detection was improved but the performance was still low because of utilizing the selective search algorithm. Finally, the Faster R-CNN suggested a region proposal network (RPN) that generated proposals based on CNN features. The proposals from RPN were then used to feed into the ROI pooling stages as in Fast R-CNN. The Faster R-CNN is considered one of the leading deep learning detection methods.

This paper presents a novel unified framework based on hybrid features that combine information from deep learning and handcrafted features to automatically detect esophageal abnormalities from endoscopic images. The CNN features are learned from the endoscopic image using a proposed DenseNet architecture and are used to generate proposals in a Faster R-CNN network. Our method integrates the DenseNet features with Gabor handcrafted features into the final detection stage of the Faster R-CNN. The contributions of this paper are shown as follows:

- We introduce a novel framework for the detection of esophageal abnormalities from endoscopic images based on the Faster R-CNN. We designed a CNN backbone network based on the DenseNet architecture to extract the CNN features.
- Gabor features are extracted from the endoscopic images and concatenated with CNN features for the ROI pooling stages in the Faster R-CNN to improve detection performance. To the best of our knowledge, it is the first-time Gabor filter responses are incorporated into the Faster R-CNN.
- The proposed model is trained end-to-end and exten-

sively evaluated on two different datasets with two types of esophageal abnormalities (Esophagitis and EAC). Our method achieved promising results on both datasets and we demonstrate that a generalized high performance can be achieved through the newly designed architecture even when using limited training data (*i.e. Miccai'15 dataset*).

This paper is structured as follows: Section II, provides an overview of the related state-of-the-art methods. Section III describes the details of the implementation of our proposed detection system. In Section IV, the dataset used in this study and evaluation metrics are described. Then in Section V the experimental results and discussion are presented. Finally, we conclude this study in Section VI.

II. RELATED WORK

In literature, methods for automatic detection of esophageal abnormalities are divided into two categories: Handcrafted features based methods and CNN based methods. This section briefly reviews methods based on HD-WLE/WLE images from both categories. More details about these methods and other techniques that utilize different examination modalities are discussed in details in [28] and [29].

- *Handcrafted Features*: Previous EAC detection methods are mostly based on handcrafted features. Sommen *et al.* [30]–[33] proposed extracting texture and color features from the original and Gabor filtered endoscopic images to detect EAC. The extracted features were classified using *Support Vector Machine* (SVM) achieving a sensitivity of 0.86 and specificity of 0.87. Additionally, the features were classified using Random Forest (RF) [34], resulting in a recall of 0.90 and precision of 0.75. Another study was proposed by Souza Jr. *et al.* [35] to evaluate the classification of EAC regions using *Speed-Up Robust Features* (SURF). The results using SVM classifier achieved a 0.89 sensitivity and 0.95 specificity on a patch-based classification. Subsequently, in [9], the Optimum-Path Forest (OPF) classifier was suggested to classify a bag-of-visual-words (BoW) designed using the SURF and Scale-Invariant Feature Transform (SIFT). The accuracy of the classifier gained efficiency of 73.8% (SURF) - 73.2% (SIFT). Later on, Souza Jr. *et al.* [36] suggested using the Color Co-occurrence Metric from a single channel as a texture descriptor of BE and EAC images. Various classifiers such as OPF, SVM and Bayesian classifiers were used for patch-based classification. The OPF achieved the best performance with an accuracy of 73.8% for (SURF) and 73.2% (SIFT).
- *CNN based methods*: Recently, CNN based methods started to draw attention for EAC detection through transfer learning. Mendel et al. [37] classified patches from HD-WLE endoscopic images into EAC or not using CNN. A 50-layer deep residual network (ResNet) [15] was constructed and learned from the ImageNet parameters to classify non-overlapping patches. The CNN model achieved a sensitivity of 0.94 and specificity of 0.88 to classify non-overlapping patches from a dataset of 100 images at a threshold 0.8. Furthermore, Reil *et al.* [38] proposed an early EAC detection using CNN transfer learning with standard classifiers (SVM and RF). Different architecture, such as *AlexNet [13], VGG'16 [14]* and *GoogleNet [39]* were evaluated with the information transferred from the non-medical domain of ImageNet using both classifiers individually. The best performance was achieved by AlexNet-SVM 0.92 area-under-the-curve (AUC) value.

Though there are various methods for esophageal abnormality detection in literature, there exists some drawback among these approaches. All the current methods investigated the detection of only one type of abnormality "*EAC*" by extracting features from non-overlapping patches/blocks within the image. However, in our method, we not only investigate the detection of *EAC* (cancerous) regions but we also examine the detection of *Esophagitis* (precancerous) regions. Another main issue in the current methods is the limited size of dataset used for training and testing the proposed methods. In our work, we train and test the model on two different datasets composed of 1000 images (Kvasir Dataset) and 100 images (Miccai'15 dataset). Furthermore, the current CNN methods mainly rely on transfer learning which means that the initial weights were learned from a non-medical domain. In our proposed model, we train the model end-to-end by learning features directly from the entire endoscopic image.

III. METHOD

In this section, we introduce our proposed esophageal abnormality detection method. The entire proposed model is shown in Fig. 2. The first step is to extract features from the input endoscopic images using the suggested DenseNet architecture. Next, the RPN generates proposals for abnormality location using the feature map generated by DenseNet. Afterward, several Gabor filter responses are extracted and concatenated with the CNN features from the DenseNet. The fused features are then used as the input to the ROI pooling layer for final classification of each proposal generated from the previous RPN stage. The implementations details of each step will be explained in the following subsections.

A. OVERVIEW OF THE FASTER R-CNN

The heart of our model is Faster R-CNN [27], which is one of the state-of-the-art object detection frameworks based on deep learning network. The Faster R-CNN is formed of two main modules. The *first module* is the RPN that is trained to propose windows for abnormal region candidates. RPN generates K possible proposals for each location using detection box called *anchor boxes* that has various sizes and ratios. There are $(W * H * K)$ possible proposals per image where W and H represent the size of the feature map output from the convolution network. The RPN network layer has two output layers; the first is a classifier layer that produces a probability if the proposed anchor box contains an object or not. The other layer is a regression layer that

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access

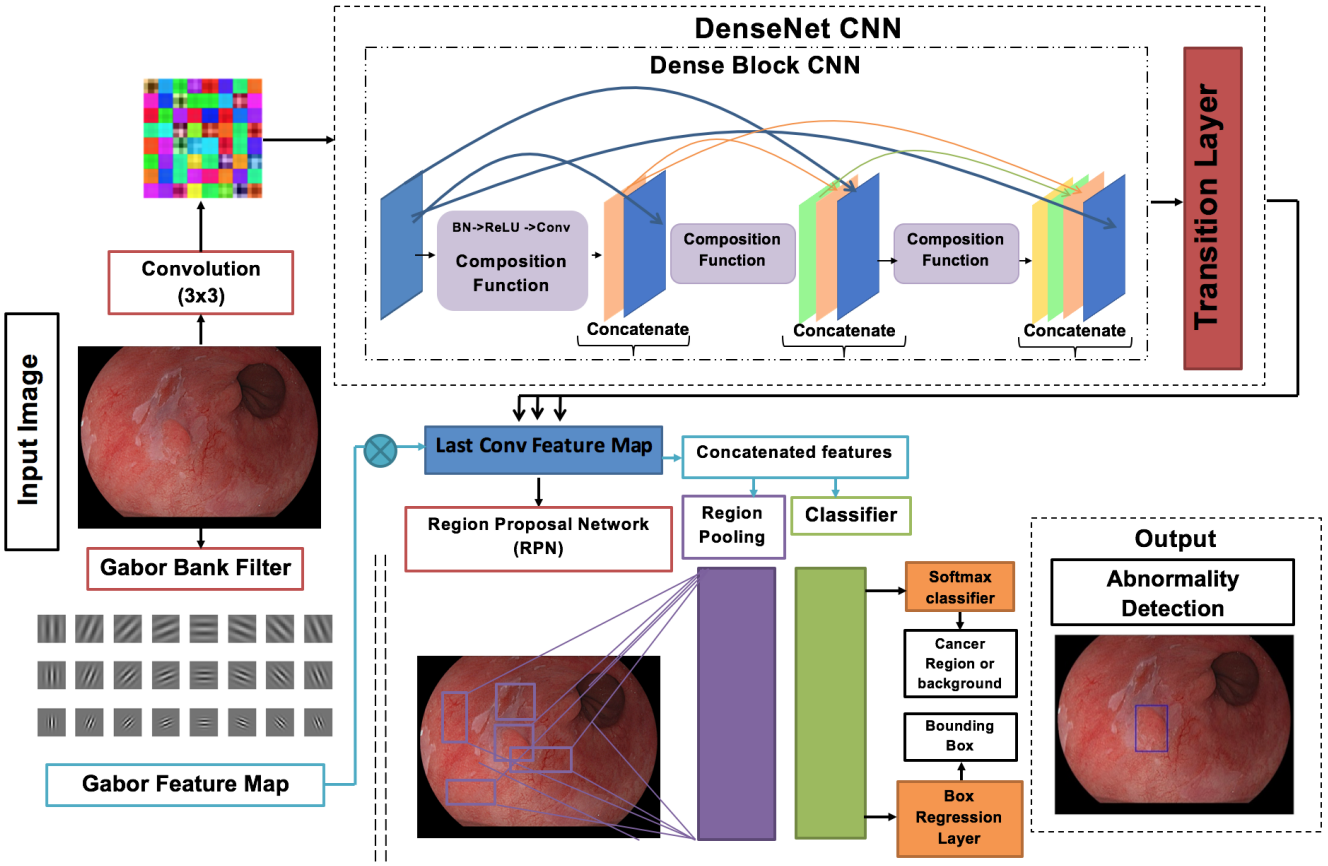


FIGURE 2: The Faster R-CNN framework outline for esophageal abnormality detection in the endoscopic images using DenseNet as a base CNN network and incorporating the Gabor features in the final detection stage. A sample of the densenet architecture with one dense block and a transition layer is illustrated as an example. The denseblock shown demonstrates the connectivity of the concatenated feature map with internal four layers.

adjusts the high probability boxes to better fit the detected object. The boxes with the highest score are called region proposals and they are sent to the next phase. During the training phase the classification and regression output from the RPN proposals rely on an Intersection-Over-Union (IoU) threshold to measure the ratio of the overlapping and union area between the ground truth and the predicted bounding box area measured as follows:

$$IoU = \frac{A_{gt} \cap A_p}{A_{gt} \cup A_p} \tag{1}$$

Here, A_{gt} is the area of the ground truth bounding box and A_p is the predicted bounding box from the regression layer.

The *second module* is the network that is trained to evaluate each proposal (abnormal candidates) from the RPN and classify the region of interest into true or a false prediction through Region-of-Interest Pooling (ROI pooling) layer. The ROI pooling reduces the size of each feature map from nominated proposal so all of them have the same size. Features in this phase are reused from the same feature map used by the RPN layer as they both share the same convolution layer. Finally, these features are used for classification. Further details about Faster R-CNN can be found in the original paper [27].

The backbone CNN network used in the original Faster-RCNN is the VGG'16 network [14], which is composed of 16 layers. It has been shown that the standard Faster R-CNN when using the VGG'16 might fail in detecting small scale objects due to information loss [40], therefore it might not be able to successfully detect the small abnormal regions with challenging appearance. In our model, we design a network architecture based on the **DenseNet** as the CNN backbone network for our Faster R-CNN model as illustrated in Fig.2.

B. DENSENET AS BASE NETWORK

DenseNets [17] has been introduced recently in literature. It reduces the connection between the input and output which helps in overcoming the vanishing gradient problem. Each layer in the DenseNet has a reduced feature map size which is important for training the CNN's on a small dataset leading to less probability of facing the over-fitting problems and ensure that there is no loss in the transmitted information [41]. Additionally, each layer receives supervision from the loss function and a regularizing effect through shorter con-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access

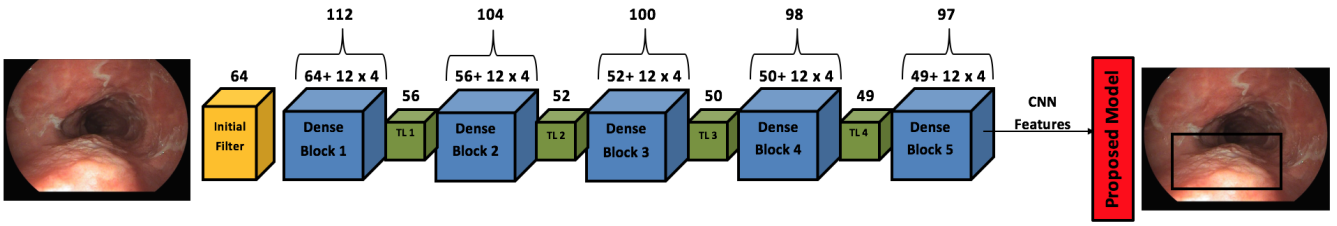


FIGURE 3: General architecture of the proposed DenseNet. An initial convolutional filter of size 64 is first performed on the input image before passing it to the first denseblock. Above each denseblock the feature map size is calculated using the number of internal layers (M) and growth rate (G). A transition layer (TL) exists between each desneblock that changes the size of the feature map.

nections leading to an easier training process. The DenseNet is mainly composed of DenseBlock, Transition Layer and Growth Rate:

- **Dense Block:**
  Each DenseNet is composed of *N* Dense Blocks. Inside each Dense Block there exists *M* layers where each layer is connected to all the consecutive layers in a feed forward manner. If x_m is denoted as the output from the m^{th} layer then it is computed as:

$$x_m = H_m([x_1, x_2, ..., x_{m-1}]) \tag{2}$$

  where H_m represents the operation of the *composite function* in this layer and a concatenation function is processed between each feature layer inside it. The concatenated features are processed through a *composite function* that consists of Batch Normalization (BN), Relu and Convolution (3x3). An example of the internal structure of denseblock that is passed on to the Transition layer is shown as a part in Fig. 2.
- **Transition Layer:**
  Between each Dense Block, a layer is introduced to decrease the spatial dimension of the features maps called *transition layer*. It is composed of Convolution (1x1) and Average Pooling (2x2).
- **Growth Rate:**
  The output from each concatenation function in (2) is *f* feature map. The size of the M^{th} layers is *f.(m-1)+*f_0, where f_0 is the number of channels of the original input image. In order to improve the parameter efficiency and control the growing of the network, the size of *f* is limited to a *growth rate G* with a small integer value. This variable helps regulate the amount of new information each layer holds.

Fig. 3 illustrates a general outline of the DenseNet with a description of the feature map size (based on $M = 4$ & $G = 12$) at each block based on the proposed implementation.

C. GABOR FEATURE

The Gabor filter is well known for texture feature representation by capturing frequency and orientation representation in the spatial domain. Generally, a gabor filter is composed of two parts *(real and imaginary)* representing the orthogonal direction. The Gabor kernel is defined as follows:

$$G(x, y, \theta_k, \lambda) = \exp\left[-\frac{1}{2}\left\{\frac{A^2_{\theta_k}}{\sigma_x^2} + \frac{B^2_{\theta_k}}{\sigma_y^2}\right\}\right] \exp\left\{i\frac{2\pi A}{\lambda}\right\} \tag{3}$$

where $A = xcos(\theta_k) + ysin(\theta_k)$, $B = -xcos(\theta_k) + ysin(\theta_k)$, λ is the wavelength and i provides the central frequency of the sinusoidal plane wave at an orientation θ_k. The orientation of $\theta_k = \frac{\pi(k-1)}{n}$ where $k = 1, 2, 3.., n$ and n demonstrates the numbers of orientations. Finally, the σ_x and σ_y denotes the standard deviations of the Gaussian envelope along the x and y axis. Fig. 4 is an example of the Gabor filter responses to endoscopic images from our dataset with different $\theta = 16$ orientations.

D. FEATURE MAP CONCATENATION FUSION

As explained earlier, to produce the output bounding box prediction, the ROI-pooling is performed on the feature map layer generated by the CNN network. In the proposed model, a Gabor feature map is generated by convolving the endoscopic image with a set of Gabor filters with different orientations. This Gabor feature map is combined with the final DenseNet feature map using concatenation fusion [42], the fused features are then used by the ROI pooling stage. The concatenation fusion takes place as:

$$F_{map} = concatenate(f_{dense}, f_{gabor}) \tag{4}$$

where, the two feature maps are stacked at the same spatial location of (i, j). Therefore, more detailed information is provided to the bounding box detection and classification from the newly concatenated feature map.

E. IMPLEMENTATION SETUP FOR EAC DETECTION

In the RPN layer of the Faster-RCNN network we adjust the anchor boxes number and sizes to the default setting as proposed in [27]. There exists k=9 anchors at each location with 3 scales (128^2, 256^2, and 512^2 pixels) and 3 aspect ratio (1:1, 1:2, and 2:1). Additionally, the loss function of the RPN stage during training process is defined as:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access

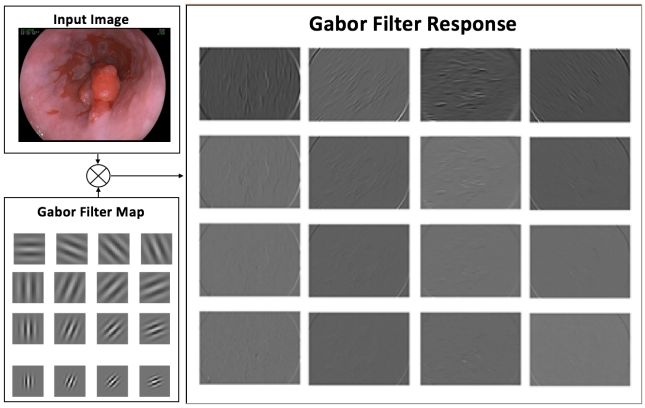


FIGURE 4: An example of Gabor Filter response with kernel size =5 with 16 different orientations.

$$L(\hat{p}_i, \hat{t}_i) = \frac{1}{N_c}\sum_i L_c(\hat{p}_i, \breve{p}_i) + \lambda \frac{1}{N_r}\sum_i \breve{p}_i L_r(\hat{t}_i, \breve{t}_i) \quad (5)$$

where, the index of an anchor is denoted by i, $\hat{p}_i$ and $\breve{p}_i$ respectively representing the predict and the ground-truth of the anchor i, being an abnormal region in the image or not. In the same manner, $\hat{t}_i$ and $\breve{t}_i$ denote the coordinates of the predicted bounding box by RPN and the ground-truth one. The total number of inputs are represented by N_c for classification layer and N_r for regression layer that is weighted by a balancing parameter λ. The L_c defines the classification loss by taking the log loss function over two classes (*abnormal candidate or not*) defined as:

$$L_c(\hat{p}_i, \breve{p}_i) = -\breve{p}_i \log \hat{p}_i - (1 - \breve{p}_i)\log(1 - \hat{p}_i) \quad (6)$$

And, L_r represent the regression loss defined as:

$$L_r(\hat{t}_i, \breve{t}_i) = L_1^{smooth}(\hat{t}_i - \breve{t}_i) \quad (7)$$

The regression loss (L_r) is only active if the ($\hat{p} = 1$) which means that the anchor boxes returned a positive candidate and it is deactivated if ($\hat{p} = 0$).

The DenseNet in our model is formed of 5 *dense blocks* with $M = 4$ internal number of layers, and a growth rate $G = 12$ that limit the network from getting too wide as the feature map will continue to grow after each dense_block. Furthermore, the transition layer applied between each *dense block* is made of (1x1) convolution layer and (2x2) average pooling layer. An initial filter of size 64 is applied to the endoscopic input image using a (3x3) convolution to create a feature map for the first denseblock (as shown in Fig. 3).

The weights are initialized randomly with a gaussian distribution ($\mu = 0$, $\sigma = 0.01$). The initial learning rate was set to 0.0003 and drops by the factor 0.1 every 1000 iteration and used a weight decay of 0.0004. The model is implemented using Keras Libary (Tensorflow backend) on a desktop with Intel Core i7 (3.6GHz processor) and an NVIDIA GeForce GTX1080 Ti with 11GB on a single GPU memory.

IV. MATERIALS AND EVALUATION METRICS

In this section, we first give details about the dataset used to evaluate the performance of the proposed model. Then the measures used in the evaluation process are described.

A. DATASET

Extensive experiments were performed to investigate the detection performance of the proposed DensNet Faster R-CNN with Gabor features on two representative datasets that include different types of esophagus abnormalities:

- *The Kvasir Dataset*:
  The Kvisar Dataset [43] is an open-access dataset that provides classified set of images inside the gastrointestinal (GI) tract. In our evaluation, we used the Esophagitis dataset that is composed of 1000 images obtained from different patients with a resolution that varies from 720×576 to 1920×1072. An expert in the field has manually annotated abnormalities in the images. Fig. 5 illustrates samples from the Kvasir dataset with the annotation by the expert.
- *EndoVis sub-challenge MICCAI'15 Dataset*:
  The dataset of the sub-challenge Early Barrett Cancer detection from *EndoVis MICCAI 2015* challenge [44] is composed of total 100 HD-WLE images with resolution of 1600×1200 gathered from 39 patients. The images are divided into 50 images without any cancer signs (Fig. 6a) obtained from 17 patients and the other 50 with cancerous regions (Fig. 6b) from 22 patients diagnosed with esophageal adenocarcinoma (EAC). Lesions found in the abnormal images have been annotated by five leading experts in the field to obtain gold standard as shown in Fig. 6c. Due to the inevitable differences between manual segmentation obtained from different experts, we took into consideration only the intersection region between the annotation from all experts for training purpose (known as sweet-spot region [45]).

Data Augmentation is introduced to the training data to increase the dataset in order to achieve better performance. It contains random rotation in different directions (45°, 135°, 225°), flipping, stretching vertically and horizontally for only 30% of the training dataset selected randomly. Therefore, the Kvasir dataset after augmentation is increased to 1900 images while the Miccai'15 dataset to 280 images. The augmented images are only included in the training phase.

B. EVALUATION MEASURES

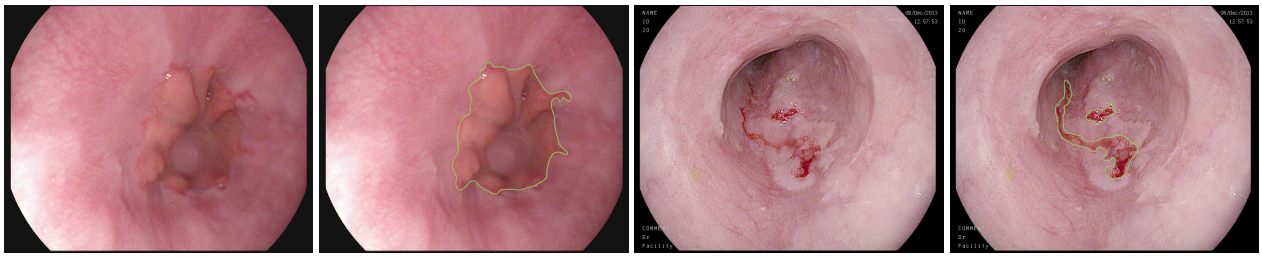
To evaluate the performance of the proposed model, the following assessment measures are employed:

$$Recall(Rec) = \frac{TP}{TP + FN} \quad (8)$$

$$Precision(Pre) = \frac{TP}{TP + FP} \quad (9)$$

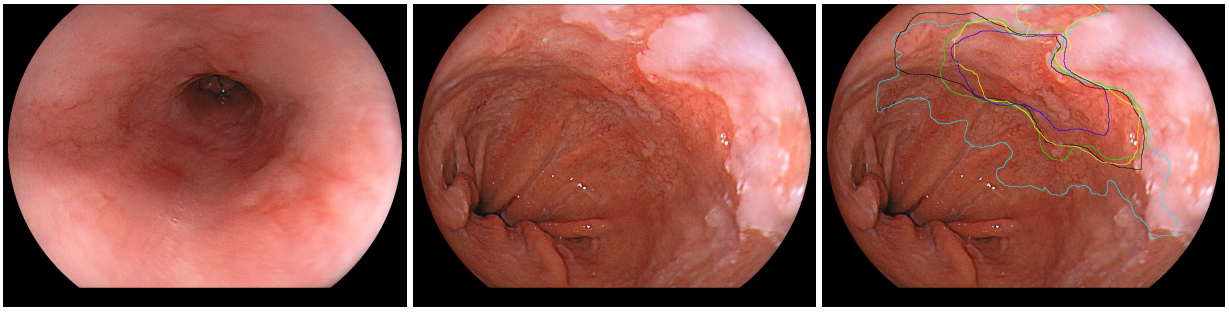
$$F1 - score = \frac{2 \times Pre \times Rec}{Pre + Rec} \quad (10)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access



(a) Esophagitis abnormality (b) Annotation by an expert (c) Esophagitis abnormality (d) Annotation by an expert

FIGURE 5: Example from the Kvasir dataset showing images with Esophagitis abnormalities (a&c) with the annotation by the expert (b&d).



(a) Normal Patient (b) EAC Patient (c) Annotation by experts

FIGURE 6: Example from the Miccai'15 dataset showing (a) Non-cancerous barrett's patient, (b) Esophageal Adenocarcinoma patient and (c) Annotation from five different experts.

where *TP (True Positive)* indicates the number of bounding-boxes that has a correct prediction in abnormal images, *TN (True Negative)* is the number of normal images that has no bounding-box, *FN (False Negative)* represent the number of abnormal images that has no prediction and *FP (False Positive)* is number of bounding boxes generated outside the abnormal ground-truth region. The bounding box is defined as a TP if it has an IoU of 0.5 or more with the ground-truth annotation and FP otherwise.

Additionally we include the following measure to evaluate the performance of detection localization by the proposed methods:

- *Mean of Average Precision (mAP)*: that measures the mean of Average Precision (AP) of the detection output. The AP measures the precision at different recall intervals where $AP = \frac{1}{11} \sum_{recall_i} Precision(Recall_i)$.

V. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, experiments are carried out to evaluate the performance of the proposed method using each dataset separately. First, experiments are conducted to investigate the effect of extracting features based on the implemented DenseNet network connection. Then, we illustrate the effect of concatenating the Gabor features with CNN features on the detection performance. Moreover, we demonstrate different visual examples of the detection output from the utilized dataset using the proposed model. Finally, we compare the performance of the method with state-of-the-art results.

A. EVALUATION OF ESOPHAGITIS DETECTION

In this section, we report the performance of our abnormality detection method in locating Esophgities regions. The Kvasir dataset was divided into 50% training, 10% validation and 40% testing by randomly selecting the images. First, to identify the effect of extracting features using DenseNet, we compare the detection results with the VGG'16 and AlexNet when used as a CNN backbone network for the Faster R-CNN. As mentioned earlier, the VGG'16 was used as the CNN backbone in the original Faster R-CNN. Table 1 displays the detection recall, precision, F1-Score and mAP values when extracting CNN feature with different CNN networks. As shown, extracting features using DenseNet improved the result of recall by 4.3% & 5.2% and precision by 2.3% & 2.6% when compared to the other two networks. This implies that utilizing the Densenet to extract features enhances the information flow throughout the network with dense connections leading to an improved performance.

Secondly, we compare the detection results after merging the Gabor features with the CNN features for the three networks. It can be seen from Table 2 that using the DenseNet with Gabor features was able to maintain the highest detection performance. Additionally, when comparing the results of Table 2 with Table 1, it can be concluded that adding the Gabor filter responses to the feature map enhances the texture information leading to an outstanding effect on the final

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access

TABLE 1: A comparison between different architectures as a backbone for the Faster R-CNN *DenseNet, VGG'16* and *AlexNet* evaluated on the Kvasir dataset.

Methods	Recall	Precision	F1-M	mAP
DenseNet	0.879	0.884	0.882	0.716
VGG'16	0.836	0.861	0.848	0.689
AlexNet	0.827	0.858	0.842	0.672

results. As shown, the results of the detection were improved from 87.9% to 90.2% in case of the DenseNet. Moreover, it had a positive impact on the other networks where the results in case of the VGG'16 were increased from 83.6% to 86.4% and 82.7% to 86.1% for the AlexNet. Furthermore, there is a 4.3% mAP improvement by the proposed model compared to using the DenseNet only which indicates a strong overall performance.

TABLE 2: A comparison of results after concatenation the Gabor features with different CNN architectures as a backbone for the Faster R-CNN evaluated on the Kvasir dataset.

Methods	Recall	Precision	F1-M	mAP
Proposed Model	0.902	0.921	0.921	0.759
VGG'16 Gabor features	0.864	0.891	0.877	0.742
AlexNet Gabor features	0.861	0.903	0.881	0.736

Moreover, we also plot the AP measure as a function of the IoU threshold in Figure 7. It can be observed that, for Esophagitis detection, the CNN network with the Gabor features outperform the network without the Gabor features. Also, our proposed model obtains a higher AP in a wide range of IoU threshold than the other methods confirming the efficiency of our designed Densenet backbone network with Gabor features in the detection process.

Furthermore, Fig. 8 provides qualitative examples of our esophagitis detection results. Figures *8a* through *8f* display samples of the images with correct detection. We find that our model is able to successfully detect various esophagitis regions of different sizes and appearances. The connection between preceding layers in DenseNet provides richer patterns, therefore, the proposed model was able to detect small regions that were not detected by the other networks such as *Fig. 8a, Fig. 8e & Fig. 8f*. Moreover, in this study, if the generated bounding box has an intersection less than a threshold of 0.5 with the ground-truth (as described earlier) we consider the bounding box a false prediction, even though it correctly detected an abnormality (i.e. if the threshold was set lower, therefore, the region would be considered a TP), Fig. *8g* & Fig. *8h* illustrate examples from these cases. Moreover, Fig. *8i* & Fig. *8j* represent samples of the incorrect prediction. Most of the false predictions made by the model capture regions that have a difference in color/texture from

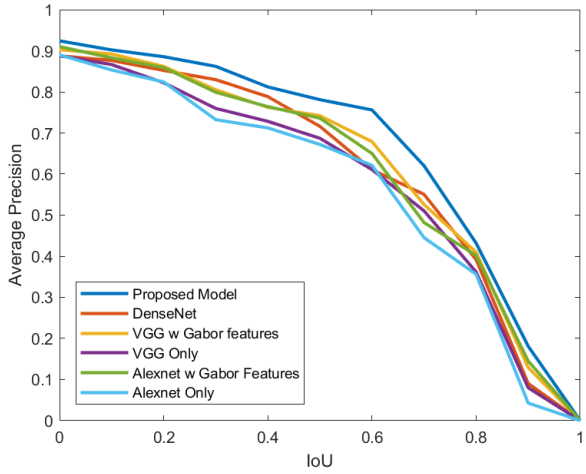


FIGURE 7: AP-IoU threshold curves using different CNN network with and with Gabor features for Esophagitis detection in Kvasir dataset.

the surrounding area. Additionally, Fig. *8k* & Fig. *8l* present negative outputs, as the detection model was not able to detect an abnormality in the endoscopic image. Overall, our model proved to have a strong performance in detecting esophagitis regions.

B. EVALUATION OF EAC DETECTION

The performance of the proposed model in detecting the EAC regions is reported in this section. For the Miccai'15 dataset, we train and validate the model on Leave-One-Patient-Out cross-validation (LOPO-CV) approach as the number of images from each patient is provided (i.e. LOPO-CV has the advantage of estimating less bias results). For the (LOPO-CV), that data is divided into N folds (*N is the number of patients*) where each fold excludes the full images of a single patient that is later used for testing and 10% of the fold is set aside for validation. First, we compare the proposed model with other CNN backbone networks for the Faster R-CNN as described in the previous section. Table 3 represents the results of the different CNN networks without Gabor features while Table 4 illustrate the results with Gabor features. From both tables, the consequences of learning features with the DenseNet is presented by increasing the accuracy of detection by 5% & 7% with Gabor features and by 2% & 4% without Gabor features when compared with VGG'16 & AlexNet respectively.. Additionally, the Gabor feature complements the feature map leading to a high recall rate in detection of the EAC region correctly with less false regions. The superior performance of the proposed model is confirmed by comparing it with the other networks. As illustrated, adding the Gabor features increased the recall from 0.90 to 0.95, the precision from 0.88 to 0.91 and F-measure from 0.89 to 0.93 when using DenseNet as the backbone network. Also, in the case of using the VGG'16 as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access

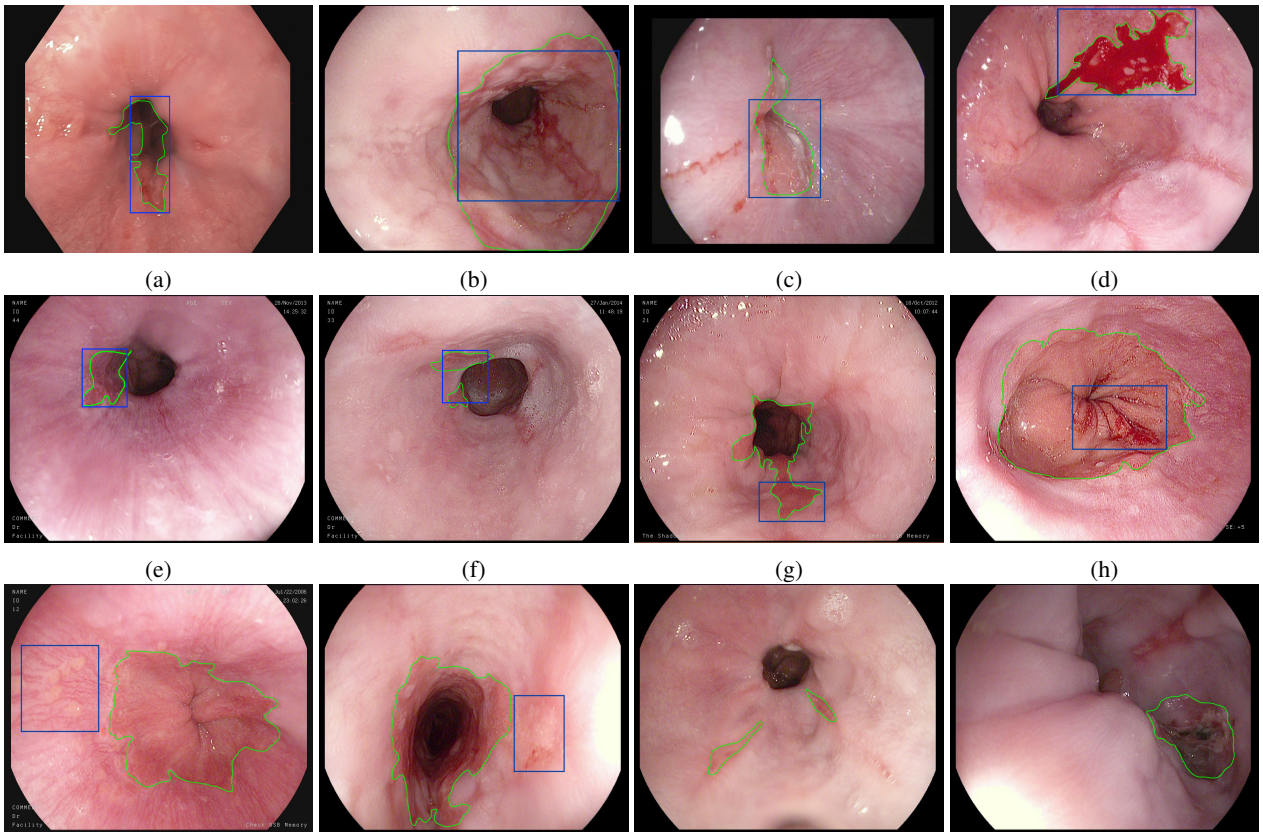


FIGURE 8: Detection examples from Kvasir dataset. The gold-standard by the expert is outlined with green in all the images. The generated bounding box by the model appears in the images with blue. From the first & second row, figures (a) to (f) represents correct detection results. Figures (g) to (j) represents samples some false predictions where (g) & (h) have an IoU<0.5 while (i) & (j) wrong location. Figures (k) & (l) shows a false negative output where the model was not able to predict any abnormality.

backbone network the recall increase from 0.88 to 0.90, the precision from 0.86 to 0.87 and F-measure from 0.87 to 0.88. And in the case of using the AlexNet as backbone network the recall increase from 0.86 to 0.88, the precision from 0.87 to 0.88 and F-measure from 0.86 to 0.88.

TABLE 3: A comparison between different architectures as a backbone for the Faster R-CNN *DenseNet, VGG'16* and *AlexNet* evaluated on the Miccai'15 dataset.

Methods	Recall	Precision	F1-M	mAP
DenseNet	0.90	0.88	0.89	0.81
VGG'16	0.88	0.86	0.87	0.78
AlexNet	0.86	0.87	0.86	0.78

Moreover, the mAP values has been increased from 0.81 to 0.84. Fig. 9 represents the AP measure as a function of the IoU threshold for the Miccai'15 dataset. As shown the proposed model achieved a high AP over different IoU thresholds compared to other networks proving the effectiveness of the model in find EAC regions.

TABLE 4: A comparison of results after concatenation the Gabor features with different CNN architectures as a backbone for the Faster R-CNN evaluated on the Miccai'15 dataset based on a LOPO-CV

Methods	Recall	Precision	F1-M	mAP
Proposed Model	**0.95**	**0.91**	**0.93**	**0.84**
VGG'16 Gabor Feature	0.90	0.87	0.88	0.82
AlexNet Gabor Feature	0.88	0.88	0.88	0.80

To visualize the output from the proposed automatic detection method, we show examples for the correctly detected lesions, false positives and missed EAC lesions in Fig. 10. As observed, the proposed method was able to successfully locate tumor regions in several EAC images, examples for correct detection with challenging cases are shown from Fig. 10a through Fig. 10d. After inspecting the missed EAC

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/ACCESS.2019.2925585, IEEE Access

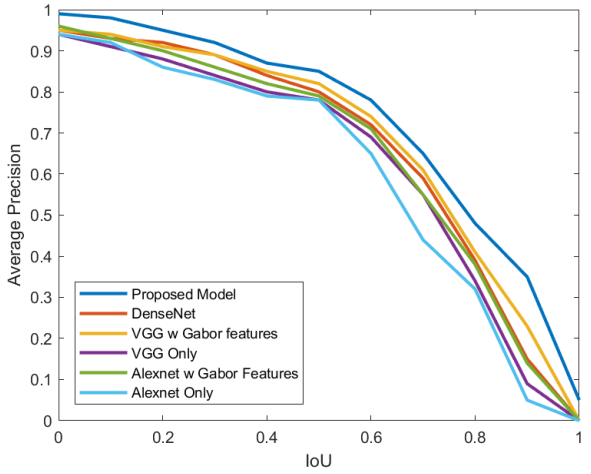


FIGURE 9: AP-IoU threshold curves using different CNN network with and with Gabor features for EAC detection in Miccai'15 dataset.

lesions, we have found that most of the missed images are the tumors that mainly have a flat surface with the esophagus (for example: Fig. 10f). The false positive in our model are mainly images with high barrett's grade or have extreme change in tissue color as shown in Fig. 10g & 10h.

C. COMPARISON WITH STATE-OF-THE-ART METHODS

To validate the effectiveness of the proposed method, we compare the results of our detection method with the results of two state-of-the-art methods reported in [30] and [37] that use the same dataset of Miccai'15 to find EAC regions. Moreover, for a fair comparison, the same validation method (LOPO-CV) is adapted. As shown in Table 5, the results of our detection methods outperformed against the state-of-the-art results regarding all evaluation measures with a *Recall: 95%, Precision: 91%, Specificity: 91%* and *F-measure: 93%*. Learning features using the proposed model achieved better results with reduced trainable parameters than [30] and [37], validating the effectiveness of reusing the features throughout the network and enhancing the model performance on limited training data.

TABLE 5: A comparison between the Proposed Model and state-of-the-art methods Sommen et al. [30] and Mendel et al. [37] on the Miccai'15 dataset based on a LOPO-CV

Methods	Recall	Precision	F1-M
Proposed Model	**0.95**	**0.91**	**0.93**
Sommen et al. [30]	0.86	0.87	0.87
Mendel et al. [37]	0.94	0.88	0.91

D. ADDITIONAL MEASURES

The differences in recall and precision calculated using the proposed model and using the DenseNet without the Gabor features were statistically evaluated for both datasets, using the paired t-test at a confidence level of 95%. The results of the two-tailed *p-value* are shown in Table 6. As shown, for the Kvasir dataset the difference between the recall and precision values for the proposed model were found to be significantly different when compared with the detection using features extracted by the DenseNet only. On the other hand, the Miccai'15 dataset showed to be significantly different only for the recall results. Moreover, the detection time during testing was also investigated. The average time to generate detection bounding boxes using our proposed model was an average of **2.34** seconds. We assume that the detection speed could be improved when using a more powerful GPU.

TABLE 6: The p-value calculated using the paired t-test to measure the difference of recall and specificity precision of proposed model on the two datasets

	Recall	Precision
Kvasir dataset	0.0055	0.00023
Miccai'15 dataset	0.0447	0.10219

VI. CONCLUSION

In this study, we present a deep learning method to automatically detect esophageal abnormalities. The Gabor filter responses calculated from endoscopic images are incorporated into the Faster R-CNN while adopting the DenseNet as the backbone network for CNN feature extraction. The dense connectivity in DenseNet improves the flow of information and the efficiency of parameters throughout the network by reusing the learned features from previous layers. The Gabor features extract local information which is fused with CNN features, therefore improving the information used by Faster R-CNN for abnormality detection. An additional advantage of the proposed method is that it is trained using the full image as an input instead of patches from the image as used by other methods [37] in the literature. Currently, in our work, we only investigated the detection of the abnormal location by using the bounding box generated by the Faster-RCNN. Future direction will include increasing the size of the dataset with more types of abnormalities such as (BE and SCC) and the investigation of segmenting abnormal regions.

REFERENCES

[1] Worldwide Cancer Data: Global cancer statistics for the most common cancers, [Online] https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data.

[2] Cancer Stat Facts: Esophageal Cancer, [Online] https://seer.cancer.gov/statfacts/html/esoph.html.

[3] Gao, Q.Y. and Fang, J.Y., 2015. Early esophageal cancer screening in China. Best Practice & Research Clinical Gastroenterology, 29(6), pp.885-893.

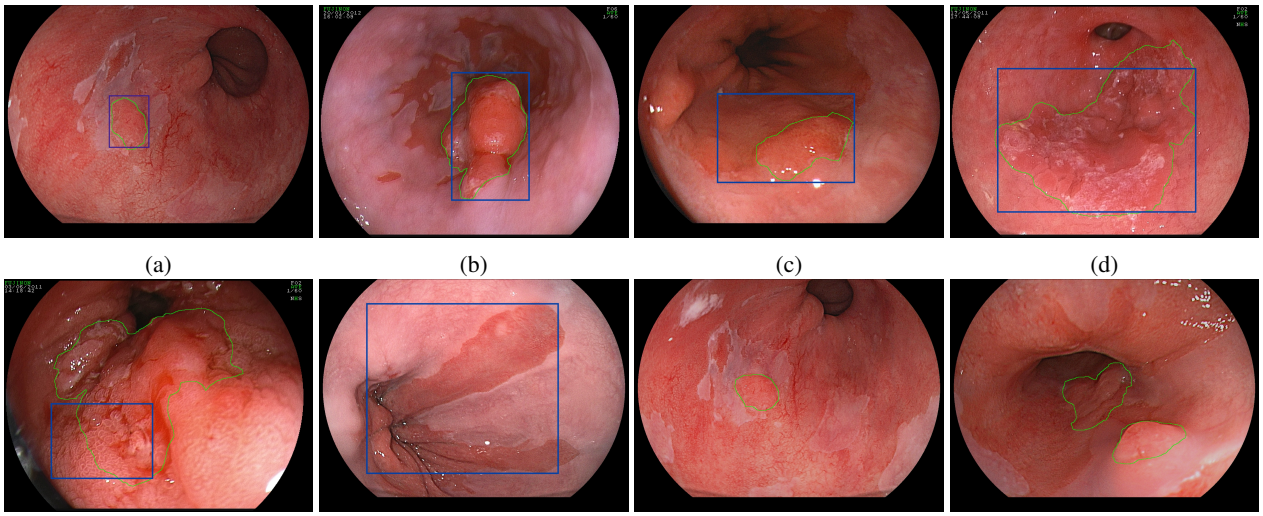


FIGURE 10: Detection examples from Miccai'15 dataset, The gold-standard of the intersection between the 5 experts (sweet-spot region) is outlined with green in all the images. The generated bounding box by the model appears in the images with blue. The first row, from (a) to (d) represents correct detection results. The first row, from (a) to (d) represents correct EAC detection results. The second-row, (e) represents a false prediction (Intersection with ground truth < 0.5 or wrong location), (f) false prediction in a non-cancerous patient and (g) & (h) both show a false negative output where the model was not able to predict any abnormality.

[4] Yui-Hsi Wang, Simon P. Hogan, Patricia C. Fulkerson, J. Pablo Abonia, Marc E. Rothenberg, "Expanding the paradigm of eosinophilic esophagitis: Mast cells and IL-9", Journal of Allergy and Clinical Immunology, Volume 131, Issue 6, 2013, pp: 1583-1585, ISSN 0091-6749, https://doi.org/10.1016/j.jaci.2013.04.010.

[5] Ghatwary, N., Ahmed, A., Ye, X. and Jalab, H., 2017, March. Automatic grade classification of Barretts Esophagus through feature enhancement. In Medical Imaging 2017: Computer-Aided Diagnosis (Vol. 10134, p. 1013433). International Society for Optics and Photonics.

[6] Menon, Shyam, and Nigel Trudgill. "How commonly is upper gastrointestinal cancer missed at endoscopy? A meta-analysis." Endoscopy international open 2.02 (2014): E46-E50.

[7] Scholvinck DW, van der Meulen K, Bergman JJ, Weusten BL. Detection of lesions in dysplastic Barrett's esophagus by community and expert endoscopists. Endoscopy. 2017 Feb;49(02):113-20.

[8] Munzer, B., Schoeffmann, K. and Boszormenyi, L., 2018. Content-based processing and analysis of endoscopic images and videos: A survey. Multimedia Tools and Applications, 77(1), pp.1323-1362.

[9] De Souza, L.A., Afonso, L.C.S., Palm, C. and Papa, J.P., 2017, October. Barrett's Esophagus Identification Using Optimum-Path Forest. In Graphics, Patterns and Images (SIBGRAPI), 2017 30th SIBGRAPI Conference on ,pp. 308-31.

[10] Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B. and SÃąnchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical image analysis, 42, pp.60-88.

[11] J. Yi, P. Wu, D. J. Hoeppner, and D. Metaxas, "Fast Neural Cell Detection Using Light-Weight SSD Neural Network", Presented at: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 860-864, Jul. 2017.

[12] Greenspan, H., Van Ginneken, B. and Summers, R.M., 2016. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Transactions on Medical Imaging, 35(5), pp.1153-1159.

[13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks". In Advances in neural information processing systems, pp. 1097-1105. 2012.

[14] K. Simonyan and A. Zisserman (2014). "Very deep convolutional networks for large-scale image recognition", Arxiv.org, 2014.[Online] Available at: https://arxiv.org/pdf/1409.1556.pdf.

[15] He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition ,pp. 770-778.

[16] Liu, W. and Zeng, K., 2018. SparseNet: A Sparse DenseNet for Image Classification. arXiv preprint arXiv:1804.05340.

[17] G. Huang, Z. Liu, L. van der Maaten and K. Weinberger, "Densely Connected Convolutional Networks", Arxiv.org, 2016. [Online] Available: https://arxiv.org/abs/1608.06993.

[18] Hosseini, S., Lee, S.H. and Cho, N.I., 2018. Feeding Hand-Crafted Features for Enhancing the Performance of Convolutional Neural Networks. arXiv preprint arXiv:1801.07848.

[19] Shi, Q., Li, W., Zhang, F., Hu, W., Sun, X. and Gao, L., 2018. Deep CNN With Multi-Scale Rotation Invariance Features for Ship Classification. IEEE Access, 6, pp.38656-38668.

[20] Luan, Shangzhen, Chen Chen, Baochang Zhang, Jungong Han, and Jianzhuang Liu. "Gabor convolutional networks." IEEE Transactions on Image Processing (2018).

[21] Yao, H., Chuyi, L., Dan, H. and Weiyu, Y., 2016, July. Gabor feature based convolutional neural network for object recognition in natural scene. In Information Science and Control Engineering (ICISCE), 2016 3rd International Conference on (pp. 386-390). IEEE.

[22] Kwolek, Bogdan. "Face detection using convolutional neural networks and Gabor filters." In International Conference on Artificial Neural Networks, pp. 551-556. Springer, Berlin, Heidelberg, 2005.

[23] Chen, Y., Zhu, L., Ghamisi, P., Jia, X., Li, G. and Tang, L., 2017. Hyperspectral images classification with Gabor filtering and convolutional neural network. IEEE Geoscience and Remote Sensing Letters, 14(12), pp.2355-2359.

[24] VilariÃśo, F., Spyridonos, P., Pujol, O., VitriÃă, J. and Radeva, P., 2006, August. Automatic detection of intestinal juices in wireless capsule video endoscopy. In null (pp. 719-722). IEEE

[25] Girshick, R., Donahue, J., Darrell, T. and Malik, J.: Region-based convolutional networks for accurate object detection and segmentation. IEEE transactions on pattern analysis and machine intelligence **38**(1), 142–158 (2016)

[26] R. Girshick, "Fast R-CNN", IEEE International Conference on Computer Vision (ICCV), 2015.

[27] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", in IEEE Transactions

on Pattern Analysis & Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 2017. DOI:10.1109/TPAMI.2016.2577031

[28] N. Ghatwary, A. Ahmed, and X. Ye. "Automated Detection of Barrett's Esophagus Using Endoscopic Images: A Survey", in *Medical Image Understanding and Analysis MIUA 2017.*, Edinburgh, UK, pp. 897-908, June 2017. Available: https://doi.org/10.1007/ 978-3-319-60964-5.

[29] L. A. de Souza Jr., C Palm, R. Mendel, C. Hook, A. Ebigbo, A. Probst, H. Messmann, S. Weber, J. P.Papa, "A survey on Barrett's esophagus analysis using machine learning", Computers in Biology and Medicine, vol. 96, pp. 203-2013 ,May, 2018. Available: https://doi.org/10.1016/j.compbiomed.2018.03.014

[30] F. Van Der Sommen, S. Zinger, E.J. Schoon and P.H.N. de With, "Supportive automatic annotation of early esophageal cancer using local gabor and color features",Neurocomputing, vol. 144, pp.92-106, Nov. 2014. DOI: 10.1016/j.neucom.2014.02.066.

[31] Van Der Sommen, F., Zinger, S. and Schoon, E.J., 2013, February. Computer-aided detection of early cancer in the esophagus using HD endoscopy images. In Medical Imaging 2013: Computer-Aided Diagnosis (Vol. 8670, p. 86700V). International Society for Optics and Photonics.

[32] van der Sommen, F., Zinger, S., Curvers, W.L., Bisschops, R., Pech, O., Weusten, B.L., Bergman, J.J. and Schoon, E.J., 2016. Computer-aided detection of early neoplastic lesions in Barrett's esophagus. Endoscopy, 48(07), pp.617-624.

[33] Setio, A.A., Van Der Sommen, F., Zinger, S., Schoon, E.J. and Peter HN de With, 2013. Evaluation and Comparison of Textural Feature Representation for the Detection of Early Stage Cancer in Endoscopy. In VISAPP (1), pp. 238-243.

[34] Janse, M.H., van der Sommen, F., Zinger, S. and Schoon, E.J., 2016, March. Early esophageal cancer detection using RF classifiers. In Medical Imaging 2016: Computer-Aided Diagnosis (Vol. 9785, p. 97851D). International Society for Optics and Photonics.

[35] Souza, L., Hook, C., Papa, J.P. and Palm, C., 2017. Barrett's Esophagus Analysis Using SURF Features. In Bildverarbeitung fÃijr die Medizin 2017 ,pp. 141-146.

[36] SOUZA JR., L. A.; EBIGBO, A.; PROBST, A.; MESSMANN, H.; PAPA, J. P.; MENDEL, R.; PALM, C. Barrett's Esophagus Identification Using Color Co-occurrence Matrices. In: CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES, 31. (SIBGRAPI), 2018, Foz do IguaÃğu, PR, Brazil. Proceedings... 2018. On-line. IBI: <8JMKD3MGPAW/3RP9Q35>. Available from: <http://urlib.net/rep/8JMKD3MGPAW/3RP9Q35>.

[37] Mendel, R., Ebigbo, A., Probst, A., Messmann, H. and Palm, C., 2017. Barrett's Esophagus Analysis Using Convolutional Neural Networks. In Bildverarbeitung fur die Medizin 2017 (pp. 80-85). Springer Vieweg, Berlin, Heidelberg.

[38] Van Riel, S., Van Der Sommen, F., Zinger, S., Schoon, E.J. and de With, P.H., 2018, October. Automatic Detection of Early Esophageal Cancer with CNNS Using Transfer Learning. In 2018 25th IEEE International Conference on Image Processing (ICIP) (pp. 1383-1387). IEEE.

[39] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Ser- manet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich, "Going deeper with convolutions" in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2015.

[40] C., Zhantao, L. Duan, G. Yang, T. Yue, Q. Chen, H. Fu, and Y. Xu., "Breast Tumor Detection in Ultrasound Images Using Deep Learning.",Springer, In International Workshop on Patch-based Techniques in Medical Imaging, pp. 121- 128., Aug. 2017.

[41] Liu, Y., Hao, P., Zhang, P., Xu, X., Wu, J. and Chen, W., 2018. Dense Convolutional Binary-Tree Networks for Lung Nodule Classification. IEEE Access, 6, pp.49080-49088.

[42] Feichtenhofer, Christoph, Axel Pinz, and Andrew Zisserman. "Convolutional two-stream network fusion for video action recognition." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1933-1941. 2016.

[43] K. Pogorelov, K.R. Randel, C.Griwodz, S. L. Eskeland, and T. de Lange, D. Johansen, C. Spampinato, D. Dang-Nguyen, M. Lux, P.T. Schmidt, M. Riegler and P. Halvorsen, "KVASIR: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection",Proceedings of the 8th ACM on Multimedia Systems Conference, Taipei, Taiwan, 2017, pp. 164-169. Available:http://doi.acm.org/10.1145/3083187.3083212

[44] Sub-Challenge Early Barrett's cancer detection, [Online] https://endovissub-barrett.grand-challenge.org

[45] F. Van Der Sommen, S. Zinger, and E.J. Schoon: "Sweet-spot training for early esophageal cancer detection". In Medical Imaging 2016: Computer-Aided Diagnosis SPIE, Vol. 9785, p. 97851B, March 2016.

NOHA GHATWARY is currently pursuing her Ph.D. degree in Computer Science with the of department Computer Science, University of Lincoln, United Kingdom. She is currently a Teaching assistant in the Computer Engineering department at Arab Academy for Science and Technology, Alexandria, Egypt. Her research interests include Medical Image Analysis, Computer Vision and Machine learning.

XUJIONG YE is a Professor of Medical Imaging & Computer Vision in the School of Computer Science, University of Lincoln, UK. Prof. Ye has over 20 years' research and development experiences in medical imaging and computer vision from both academia and industry. Her main research is to develop computational models using advanced medical image analysis, computer vision and artificial Intelligence to support clinicians in decision-making.

MASSOUD ZOLGHARNI received his PhD degree in Biomedical Engineering from Swansea University, UK in 2010. He is currently an Assistant Professor at the School of Computing and Engineering, University of West London. He is also a Research Associate in the National Heart and Lung Institute, Imperial College London. His research interests include Computer Vision, Medical Imaging, Machine Learning, and Numerical Simulations.